

영역 기반 영상 검색을 위한 다중클래스 피드백 알고리즘

고 병 철[†] · 남 재 열^{**}

요 약

본 논문에서는 영역기반 영상검색의 성능 향상을 위한 피드백 알고리즘으로 다중 클래스를 갖는 확률적 신경망(Probabilistic Neural Networks)을 이용한 방법론을 제안하고 이를 영역기반 영상 검색 시스템인 FRIP(Finding Regions In the Pictures) 시스템에 적용하였다. 본 논문에서 제안하는 피드백 알고리즘은 특징 벡터가 독립적이라는 가정을 할 필요가 없으며 보다 상세한 분류를 위해 추가적인 클래스들을 추가할 수 있도록 허용하고 있다. 또한 단지 4개 층(layer)만을 가지고 있으므로 학습을 위한 계산시간이 적게 든다는 장점이 있다. 추가적으로 다음단계에서의 성능 향상을 위해 분류 단계에서 사용자의 이전 피드백 행동을 모두 히스토리(history)로 모두 기억시켜 놓고 다음 단계를 위한 가중치 학습을 위해 사용하도록 한다. 히스토리를 사용함으로써 제안하는 알고리즘은 사용자의 주관적 의도를 보다 정확하게 파악 할 수 있을 뿐만 아니라 학습을 위해 이전 단계만을 사용 했을 때 발생할 수 있는 성능 감소를 막을 수 있다. 본 논문에서는 Corel-photo CD에서 3000장의 자연 영상을 무작위로 추출하여 기존의 방법론들과 제안하는 방법론의 성능을 측정하여 본 논문에서 제안하는 방법론이 성능이 우수함을 증명하였다.

키워드 : 영역기반 영상 검색, 피드백 학습, FRIP, 확률적 신경망, 피드백 히스토리

Multi-class Feedback Algorithm for Region-based Image Retrieval

Byoung-Chul Ko[†] · Jae-Yeal Nam^{**}

ABSTRACT

In this paper, we propose a new relevance feedback algorithm using Probabilistic Neural Networks(PNN) while supporting multi-class learning. Then, to validate the effectiveness of our feedback approach, we incorporate the proposed algorithm into our region-based image retrieval tool, FRIP(Finding Regions In the Pictures). In our feedback approach, there is no need to assume that feature vectors are independent, and as well as it allows the system to insert additional classes for detail classification. In addition, it does not have a long computation time for training because it only has four layers. In the PNN classification process, we store the user's entire past feedback actions as a history in order to improve performance for future iterations. By using a history, our approach can capture the user's subjective intension more precisely and prevent retrieval performance errors which originate from fluctuating or degrading in the next iteration. The efficacy of our method is validated using a set of 3000 images derived from a Corel-photo CD.

Key Words : Region-Based Image Retrieval, Feedback Learning, Frip, PNN, Feedback History

1. 서 론

최근 대부분의 내용기반 영상/비디오 검색 시스템에서 연관성 피드백(relevance feedback) 알고리즘은 가장 많이 사용되는 검색 모델 중의 하나이다. 이러한 모델은 일반적으로 사용자와의 상호작용을 통해 검색 성능을 향상시키는데 이용되고 있으며 또한 사용자의 선택에 의해 질의 결과를 점차적으로 개선하도록 하는데 목적이 있다. 연관성 피드백을 사용하지 않는 검색 시스템들[1-3]의 경우 사용자의 주

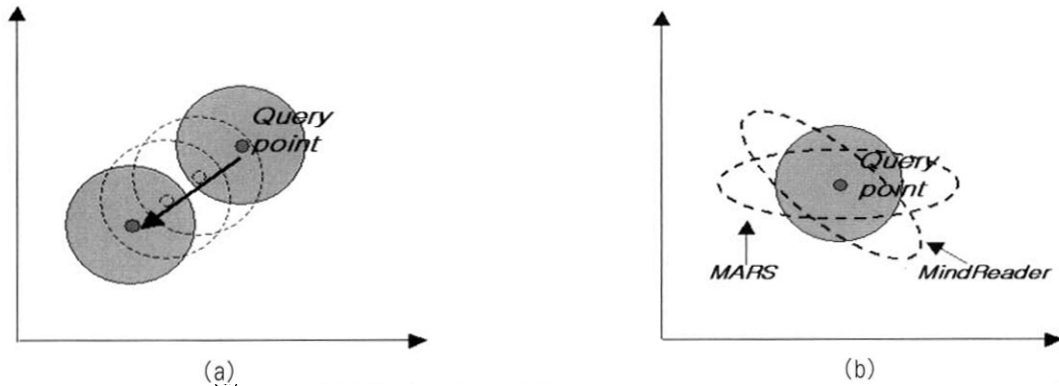
관적인 인식뿐만 아니라 고 수준(자동차, 배, 호랑이 등)의 개념들을 효과적으로 검색 시스템에 반영 하지 못하는 단점이 있다. 예를 들어 만약 사용자가 '빨간 자동차'를 검색하고자 할 경우 검색 시스템이 단지 저 수준 특징(색상, 질감, 모양 등)만을 사용한다면 빨간 색상을 가지는 사각형 물체만을 검색 결과로 보여 줄 것이다. 이 경우 사용자가 검색 결과에 만족하지 못하더라도 손으로 특징 벡터의 가중치를 일일이 조절하거나 새로운 영상을 질의하는 것 이외에는 검색 결과를 갱신할 수 있는 방법이 없다. 하지만 연관성 피드백 방법론[4-6]의 경우, 검색은 사람과 검색 시스템 사이에 상호작용적으로 진행 된다. 연관성 피드백의 주요한 기능은 사용자의 피드백을 기반으로 고 수준의 개념들과 저

* 본 연구는 산업자원부 지방기술혁신사업(RT104-01-01) 지원으로 수행되었음.

† 정 회 원 : 계명대학교 컴퓨터공학과 겸임강사

** 정 회 원 : 계명대학교 컴퓨터공학과 부교수

논문접수 : 2006년 3월 31일, 심사완료 : 2006년 6월 9일



(그림 1) 세 가지 알고리즘에 대한 특징 공간에서의 표현, (a) 질의 점 이동 알고리즘 : 균일한 거리를 갖는 원형 형태의 가중치 모델, (b) 가중치 재계산 알고리즘 : MARS-특징 공간 축에 평행한 타원 형태의 가중치 모델, MindReader-특징 공간 축에 평행하지 않은 일반화된 타원 형태의 가중치 모델

수준의 특징들 사이에 연결을 설정하는 것이다[6].

일반적으로 피드백 시스템에서는 사용자는 단지 질의를 진행시키기 위해 초기 검색된 영상에 대해 유사 영상과 비 유사 영상을 표시하기만 하면 된다. 이러한 사용자의 피드백 동작에 의해 유사성 추정을 위한 특징 가중치가 다시 재조정되게 되거나 질의 점(query point)이 이동되게 된다.

본 논문에서는 유사 혹은 비 유사만을 구분하여 가중치를 갱신하는 이진 클래스 방식의 기존 방법에서 벗어나 다중 클래스 연관성 피드백 알고리즘을 제안하고 이를 기존에 제안한 영역기반 영상검색 도구인 FRIP(Finding Regions In the Pictures) 시스템[9-10]에 적용하였다. 본 시스템의 피드백을 위한 사용자 메뉴는 매우 유사(VS: Very Similar), 유사(S: Similar), 비 유사(D: Dissimilar), 매우 비 유사(VD: Very Dissimilar)의 4 개의 클래스로 구성된다. 사용자가 선택한 예제 영역에 의해 유사한 영역을 포함하는 초기 검색 영상들이 결과로 제시되고 사용자가 4개의 클래스 중 가장 유사하다고 판단되는 클래스를 선택하는 피드백 동작이 완료된 후에 각 특징을 위한 가중치는 사용자의 주관성을 반영하는 방향으로 재조정되고 이 가중치는 다음 단계의 유사성 추정을 위해 사용된다.

본 논문의 구성은 다음과 같다. 2장에서는 연관성 피드백 알고리즘의 기존 관련 연구에 대해서 살펴보고 3장에서는 다중 클래스 연관성 피드백 알고리즘이 포함된 FRIP 시스템의 전체적인 내용을 살펴본다. 4장에서는 PNN을 이용한 다중 클래스 피드백 알고리즘에 대해 설명한다. 5장에서는 실제 자연영상을 사용하여 본 논문에서 제안하는 방법론의 성능을 비교해 보고 마지막으로 6장에서 본 논문에 대한 결론을 제시한다.

2. 기존의 연구

기존의 연관성 피드백 알고리즘은 크게 질의 점 이동 방식(query-point movement)과 재 가중치 계산(re-weighting)의 두 가지 방법론으로 분류 된다[8].

질의 점 이동 방식은 (그림 1)의 (a)와 같이 질의 점을

사용자가 선택한 ‘유사 영상 집합’으로 향하게 하고 ‘비 유사 영상 집합’으로 부터는 멀어지게 함으로써 이상적인 질의 점을 예측하도록 하는 방식이다. 질의 점 이동 방식을 위한 Rocchio의 공식은 수식 (1)과 같다.

$$z' = \alpha z + \beta \frac{1}{n_r} \sum_{X \in R_r} X - \gamma \frac{1}{n_i} \sum_{X \in R_i} X \quad (1)$$

수식(1)에서 z' 은 새로운 질의 위치를 의미하고, z 는 초기 질의 위치를 의미한다. R_r 은 유사 영상 집합을 의미하고 R_i 는 비 유사 영상 집합을 의미한다. n_r 은 유사 영상 집합의 총 원소 수를 의미하며 n_i 는 비 유사 영상 집합의 총 원소 수를 의미한다. 세 개의 파라메타 α, β, γ 는 ad-hoc 파라메타로 시스템에 따라 사용자가 수치 값을 조절해 줄 수 있고 가중치의 합은 1이다.

가중치 재계산(re-weighting)[6] 방법은 MARS 시스템에서 처음으로 제안되었다. 가중치 재계산 방법에서는 하나의 특징 공간 축 상에 존재하는 유사 영상들의 특징 값에 대한 분산이 크다면 데이터들 사이에 연관성이 적은 것으로 해당 특징 공간 축 상에서 낮은 가중치 w_j 를 부여한다. 반대로 낮은 분산은 연관된 영상들 사이에 특징 값이 유사한 것을 의미함으로 상대적으로 높은 가중치를 부여 한다. 분산은 질의 점과 예제 영상들의 특징 벡터들 간의 유클리디언 거리로 계산된다. 가중치 결정은 연산식 $w_j = 1/\sigma_j$ 에 의해 분산에 반비례하는 가중치를 갖게 된다. MARS에서 사용한 가중치 재계산 방법은 좋은 성능을 보여주지만 이론적 기반이 약하고 ad-hoc 휴리스틱에 기초한 방법이라는 한계를 가지고 있다. (그림 1)의 (b)는 MARS에서 사용한 특징 공간에서의 가중치 모델을 보여주는 것으로 특징 벡터가 2차원 공간일 경우 질의 점과 각 예제 영상들간의 가중치 모델은 항상 특징 공간의 축 방향과 일치한다. 이것은 결국 한쪽 특징 벡터의 가중치가 크다면 그만큼 상대적으로 다른 특징 벡터의 가중치는 작아짐을 의미한다. MARS의 이러한 문제점을 해결하기 위해 MindReader[8]에서는 ad-hoc 휴리스틱을 사용하지 않고 가중치 학습에 대한 수학적 프레임워

크를 개발하였다. MindReader에서는 질의 점으로부터 예제 영상 사이에 가중치 유클리디언(weighted-Euclidean) 거리를 구함으로써 특징 공간 축에 평행하지 않고 회전된 가중치 모델(점선)을 얻을 수 있다. 따라서 이러한 특징 벡터에 대한 가중치 결정은 사용자의 주관적 인식이 더 정확하게 검색 과정에 반영될 수 있다는 장점이 있다. 하지만, 예제 영상이 많고 특징 벡터의 차원이 클 경우 각 영상에 대해 모두 유클리디언 거리 계산을 위한 가중치 공분산을 계산해야 함으로 계산 시간이 많이 걸리는 단점이 있다.

두 가지 주요 방법 이외에, 최근 몇 년 동안 다양한 형태의 피드백 알고리즘이 제안되었는데 Peng[5]은 확률적 연산에 의한 연관성 피드백 알고리즘을 제안하였고, Vasconcelos[11]와 Meilhac[12]은 베이시안 추론을 통해 가중치를 결정하는 방법을 제안 하였다. 이 밖에 Macarthur[13]는 결정 트리(decision tree)를 이용한 방법을 제안하였고, Hong[14]은 SVM(Support Vector Machine)에서 결정면(hyper-plane)과 서포트 벡터 사이의 거리를 이용한 가중치 결정 방법을 제안 하였다. 하지만 대부분의 방법은 사용자에게 유사 혹은 비유사의 두 가지 클래스에 대해서만 선택하도록 하고 이를 이용하여 가중치를 갱신한다. 가중치는 다시 유사성 측정에 사용되는데 사용자가 선택하기에 애매모호 하거나 불확실한 영상이 포함될 경우 오히려 가중치 결정에 나쁜 영향을 미칠 수 있다.

본 논문에서는 확률적 신경망(PNN: Probabilistic Neural Networks)을 이용한 새로운 연관성 피드백 알고리즘을 제안하고 이를 영역기반 영상 검색 시스템인 FRIP(Finding Regions In the Pictures)에 적용하였다. 본 논문에서 제안하는 PNN 기반 방법의 경우 사용자 선택의 범위가 단순히 유사, 비유사의 두 클래스가 아닌 매우 유사(VS: Very Similar), 유사(S: Similar), 비 유사(D: Dissimilar), 매우 비 유사(VD: Very Dissimilar)의 4가지 멀티 클래스 분류가 가능하다. 따라서 본 논문에서는 VS, S, D, VD의 4가지 클래스를 통해 가중치를 갱신하고 이를 검색에 사용하여 성능을 향상 시키도록 하였다. 또한, 히스토리(history) 개념을 추가하여 이전 단계에서 학습한 내용을 기억 시켜 놓음으로써 이전 단계에서 발생한 오류가 다음단계에서 발생하지 않고 검색 성능을 높일 수 있도록 하였다.

3. FRIP 시스템 개요

FRIP 시스템[9]은 영역을 분할하고 유사성 측정을 위해 각 영역으로부터 5개의 특징 값을 추출하여 검색에 사용하는 영역기반 영상 검색 시스템이다. 영역 분할을 위해서는 베이시안 알고리즘과 질감 확률 분포를 이용한 세 개의 원형 필터를 사용한다. 첫 번째 단계에서 세 개의 적응적 원형 필터를 영상 안에 포함된 질감 정보의 양에 따라 다르게 적용하고 두 번째 단계에서 첫 번째 단계의 결과에서 나온 작은 패치(patch)들을 영역 병합과 영역 레이블링(labeling)을 통해 임계값 이하 까지 반복하는 과정을 실시한다. 최종적으로 영역의 개수가 일정 이하인 작은 영역들을 인접한

최대 크기 영역으로 병합 하는 과정을 적용한다. 영역이 분할된 후에 유사성 측정을 위해 다음과 같은 5개의 특징과 유사성 측정을 위한 거리 측정법을 사용한다.

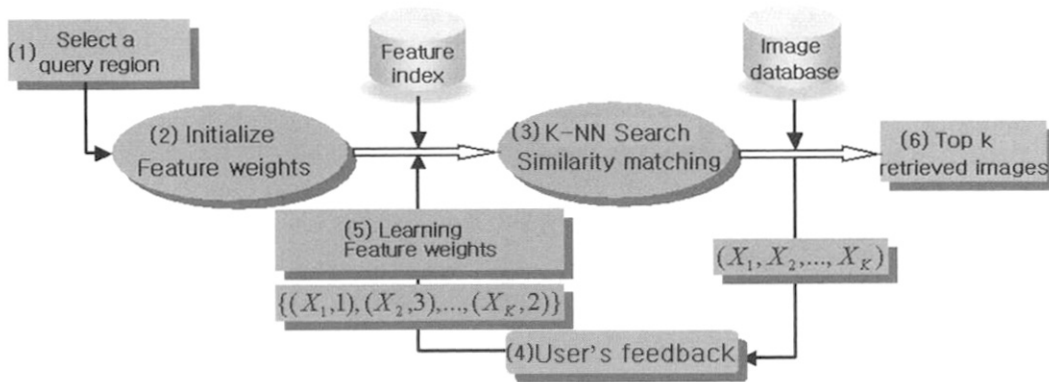
- 색상 정보: CIE-LAB 색상 공간으로부터 각 영역의 L, a*, b* 색상의 평균 및 분산 추출 한다. 질의(q) 및 데이터베이스 영역 (t) 간의 색상 유사성 측정을 위해 city-block 거리 측정법을 사용하고 이를 $d_{q,t}^c$ 라 표시한다.
- 질감 정보: 질감 성분을 위해 웨이블릿 변환된 각 영역의 수직 및 수평 방향의 웨이블릿 계수의 크기를 추출한다. 두 영역사이의 거리 측정을 위해 city-block 거리 측정법을 사용하고 이를 $d_{q,t}^g$ 라 표시한다.
- 위치 정보: 위치 성분을 위해 각 영역의 중심점의 좌표를 측정한다. 두 영역 사이의 거리 측정을 위해 유클리디언 거리 측정법이 사용되며 이를 $d_{q,t}^l$ 라 표시한다.
- 크기 정보: 영역의 크기 성분을 추출하기 위해 각 영역의 화소 수를 전체 영상의 크기로 나눈 값을 사용한다. 두 영역사이의 거리 측정은 city-block 거리 측정법을 사용하고 이를 $d_{q,t}^s$ 라 표시한다.
- 모양 정보: 효과적인 모양 유사성 측정을 위해 MRS (Modified Radius-based Signature)를 제안하고 이를 $d_{q,t}^m$ 라 표시한다.

5개의 특징의 특징 정보와 거리 측정법을 이용하여 질의 영역과 데이터베이스 영상 안에 포함된 영역들간의 최종 유사성 측정은 각 측정값들의 선형 조합을 통해 이루어지며 각 측정값들의 가중치는 사용자가 인위적으로 슬라이드 바를 이용하여 조절하지 않는다면 모두 1의 값을 갖는다. FRIP 시스템에 대한 자세한 설명은[9]에 언급되어 있다.

본 논문에서는 새로운 연관성 피드백 알고리즘을 제안하고 이를 FRIP 시스템에 결합하였다. 결합된 시스템의 동작 과정을 살펴보면, 우선 사용자가 질의 영상을 선택하고 '영상 분할' 버튼을 눌러 영상을 영역단위로 분할한다. 그런 뒤에 검색 결과로 보여질 상위 k를 선택하고 검색하고자 하는 영역을 클릭한다. 사용자는 보다 정확한 검색을 위해 영상의 특성에 따라 5개의 특징을 선별적으로 선택이 가능하다. 검색 후 k개의 검색 영상이 결과로 나타나면 사용자는 자신의 주관에 따라 결과 영상을 매우 유사(VS), 유사(S), 비 유사(D), 매우 비 유사(VD)의 하나를 선택한다. 만약 판단하기 어려울 경우 선택하지 않아도 된다. 피드백 동작이 완료되면 각각의 특징을 위한 가중치가 사용자의 주관 유사한 방향으로 제조정 된다. 피드백 동작에 대한 과정은 (그림 2)에 간략하게 설명되어 있다.

4. 다중 클래스 연관성 피드백을 위한 확률적 신경망

신경망은 학습 데이터를 기반으로 하는 패턴분류에 자주 사용되는 방법이다. 여러 가지의 신경망 중 오류 역전과



(그림 2) 전체적인 피드백 동작 과정(번호는 검색 과정의 순서를 나타냄)

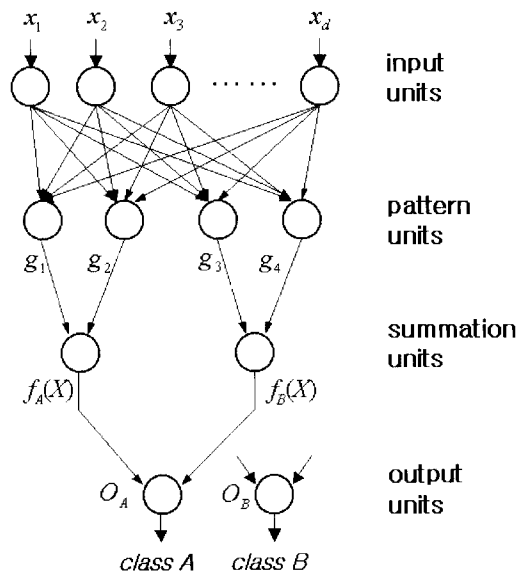
(Back-propagation) 학습기는 신경망에서 가장 많이 사용되는 방법이지만 학습을 위한 많은 계산 시간을 요구하며 지역적 국소(local minima)에 빠질 수 있는 단점이 있다[15]. 본 논문에서는 사용자 피드백 학습을 위해 오류 역전파 방법을 사용하지 않고 확률적 신경망(PNN: Probabilistic Neural Network)을 변형하여 새로운 연관성 피드백 방법론을 제안한다. 확률적 신경망의 경우 특징 벡터가 독립적이라는 가정을 할 필요가 없고 시스템으로 하여금 보다 상세한 분류를 위해 클래스를 추가할 수 있도록 허용한다. 또한, 단지 4개의 층(layer)만을 가지고 있기 때문에 학습을 위한 계산 시간이 짧다는 장점이 있다.

4.1 확률적 신경망

확률적 신경망의 망구조는 오류 역전파 알고리즘과 유사하다. 가장 중요한 차이점은 시그모이드 활성화 함수가 exponential을 포함하는 한 개의 클래스 함수로 대체된다는 점이다. 확률적 신경망의 중요한 장점은 다음과 같다.

- 학습은 단지 한 번의 과정(pass)만을 거친다.
- 결정면의 모양은 필요에 따라 적절한 스무딩 파라미터를 선택함에 따라 복잡하거나 또는 단순하게 설계할 수 있다.
- 필요에 따라 추가적인 클래스 삽입이 가능하다.

확률적 신경망의 중요한 단점은 모든 샘플들이 반드시 저장되어 있어야 하고 새로운 분류를 위해 사용되어야 한다는 점이지만 최근에는 메모리의 가격이 저렴해지고 고용량의 메모리로 인해 이러한 문제점은 확률적 신경망의 장점에 비해 문제점이 되지 못한다[16-17]. (그림 3)은 확률적 신경망의 구조를 보여주고 있다. 확률적 신경망은 d 개의 입력 유닛(input unit)과 n 개의 패턴 유닛(pattern unit), s 개의 합 유닛(summation unit), c 개의 출력 유닛(output unit)으로 구성되어 있다. 입력 층을 포함하고 있는 입력 유닛은 각각의 n 개의 패턴 유닛과 연결되어 있다. 각각의 패턴 유닛은 입력 벡터 X 와 패턴 유닛의 가중치 벡터 W_k 의 내적(dot product)으로 구성되고 (수식 (3)) 합 유닛으로 결과를 전송하기 전에 net_k 에 대해 비선형 연산을 수행한다. 합 유닛은 패턴 유닛



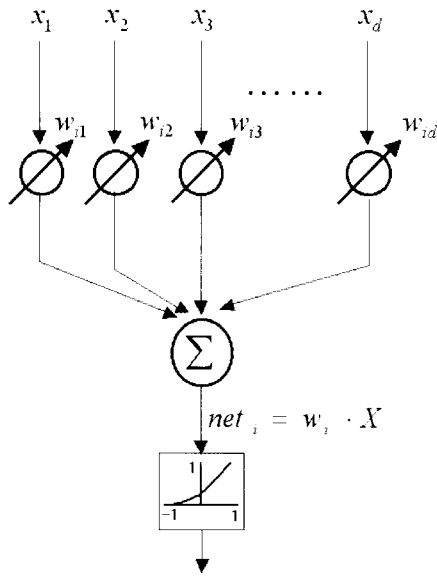
(그림 3) 4개의 패턴 유닛과 2개의 출력 유닛을 가지는 확률적 신경망의 구조

으로부터 받은 값을 합하는 역할만을 한다. 최종적으로 합 유닛은 출력 유닛 중의 한 유닛으로 연결된다. (그림 3)에서 보는 것과 같이 확률적 신경망은 4개의 층으로 구성되어 있으며 필요에 따라 패턴 유닛의 수와 출력 유닛의 수를 조절할 수 있다.

$$net_k = X \cdot W_k \tag{3}$$

(그림 4)는 확률적 신경망에서 패턴 유닛의 구조를 보여주는 것으로 하나의 패턴 유닛은 입력 벡터와 패턴 유닛의 가중치 벡터를 내적하고 그 값을 시그모이드 함수에 입력시켜 얻을 수 있다.

본 논문에서 제안하는 확률적 신경망의 훈련 과정은 다음과 같다. 우선 각각의 훈련 집합(본 논문에서는 사용자가 선택한 영상)의 패턴 x 를 $[0 \leq 1 - x_i \leq 1]$ 범위로 정규화 한다. 본 논문에서는 특징 벡터의 차원을 줄이기 위해 입력 벡터 X 를 질의 영역과 검색 결과에서 사용자가 선택한 각 영역간의 특징 값의 거리(distance)로 정하였다. 즉, 입력 패턴 x_1 은 색상 거리 값을 의미하며 5개의 특징을 사용했을 경우



$g_i = \exp[(net_i - 1) / \sigma^2]$
 (그림 4) 확률적 신경망의 패턴 유닛 구조

입력 벡터의 차원은 $x_1 \sim x_5$ 까지 5가 된다. 질의 영역과 검출된 결과 중에서 사용자가 분류한 훈련 영상간의 특징 거리가 클수록 입력 벡터는 0~1 사이의 작은 값을 갖게 된다. 이렇게 정규화된 훈련 패턴은 입력 유닛에 위치한다. 초기에 입력 유닛과 패턴 유닛을 연결하는 가중치는 $w_i = x_i$ 과 같이 설정된다. 그런 뒤에 활성화함수에 의해 패턴 유닛의 출력 값이 계산되면 이 값은 모든 합 유닛에 전달되는 것이 아니라 각각의 패턴 유닛에 대응되는 합 유닛으로 전달되고 최종적으로 출력유닛으로 전달된다. 이러한 과정은 각각의 남은 훈련 영상을 통해 반복해서 이루어지고 $w_i = x_i$ 와 같이 연속적인 패턴 유닛의 가중치를 결정한다. 여기서 j는 패턴 유닛의 수를 의미하는 것으로 본 논문에서는 VS, S, D, VD 와 같이 네 개의 패턴을 사용했으므로 j는 4가 된다. 훈련 과정이 끝난 후에 망은 입력 유닛과 출력 유닛 사이에 완전한 연결이 형성된다.

본 논문에서는 1장에서 설명한 것과 같이 사용자에게 VS, S, D, VD와 같이 4개의 연관성 피드백을 지원한다. 따라서 각 패턴 유닛은 총 4개가 된다. 훈련 과정에 대한 설명은 다음과 같다.

```

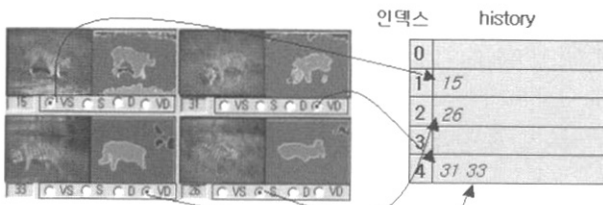
Initialize j=0, j=1,...,4; i=1,...,5
do
  i=i+1
  do{
    j=j+1
     $\bar{x}_{ji} = E[j|x_{ji}] = \sum_{n=1}^k x_{ji} \cdot 1(S_n \equiv j) / \sum_{n=1}^k 1(S_n \equiv j)$ 
     $x_{ji} = \exp(\bar{x}_{ji}) / \sum_{m=1}^5 \exp(\bar{x}_{jm})$  (정규화)
     $w_{ji} = x_{ji}$  (가중치 결정)
  }while(j < 4)
}while(i < 5)
    
```

식(4)~(6)에서 j는 4개의 패턴 유닛(VS, S, D, VD)을 나타내고 i는 5개의 특징 벡터를 나타낸다. S_n 은 k개의 결과 영상에서 사용자에게 의해 선택된 패턴을 의미한다. $\bar{x}_{ji}(=E[j|x_{ji}])$ 는 k개의 검색 결과 영상 중 사용자에게 의해 선택된 j 패턴으로부터 계산된 i번째 특징 벡터 x_{ji} 의 기대치를 의미한다. 수식 (4)에서 1(·)은 지시(indicator) 함수를 나타낸다. 즉, 만약 패턴 유닛 번호가 사용자에게 의해 선택된 패턴 번호와 일치 할 경우 1(·)은 1을 리턴 하고 그렇지 않을 경우 0을 리턴 한다. 결과적으로 수식 (4)의 분모는 k개의 S_n 중에서 j 패턴과 같은 영상의 총 수를 의미하며, 분자는 그 중 i 번째 특징 벡터의 값을 의미한다. 수식 (5)에 의해 j 패턴에서 모든 특징 벡터의 평균값의 합으로 각각의 특징 벡터 값을 나누는 정규화 과정을 실시하고 정규화 후에 실제 특징 벡터의 가중치는 특징 벡터의 출력 x_{ji} 이 된다. 수식 (5)에서 exponential 형태의 가중치 부여 방식은 지역적 특징 연관성의 변화에 민감하고 선형적 방식 보다 좋은 성능을 보여 줌으로[4-5], 본 논문에서는 이 방법을 적용하여 정규화 과정을 실시하였다.

4.2 연관성 추정(Relevance Estimation)

패턴 유닛의 가중치 벡터 w_k 가 결정되어 망에 대한 학습이 끝난 후에, 정규화된 테스트 패턴(1-X)가 입력 유닛에 위치한다. 여기서 '1'은 모두 1의 값을 갖는 1x5 벡터를 나타낸다. 각각의 패턴 유닛은 수식 (7)에 의해 망 활성화(net activation) 연산식 net_k 를 계산한다. 본 논문에서는 입력 벡터가 (1-X)의 형태를 가짐으로 망 활성화 연산식은 $w_k' \cdot (1-X)$ 와 같이 변형된다. 그런 다음 합 유닛으로 활성화 단계의 값을 출력하기 전에 오류 역전파에서 주로 사용되는 시그모이드 활성화 함수 대신에 net_k 에 대해 수식 (8)의 $\exp[(net_k - 1)/\sigma^2]$ 의 비선형 방정식을 적용한다. 여기서 σ 는 사용자에게 의해 조절 될 수 있는 매개변수로 가우시안 원도우의 폭을 결정해 준다. 패턴 유닛의 출력은 수식 (10)에 의해 대응되는 합 유닛으로 연결된다. 분리된 각 패턴 유닛은 매번의 패턴 학습을 위해 필요하다. (그림 3)이 보여주는 것과 같이 같은 패턴 유닛(VS, S와 D, VD)은 다른 합 유닛에 의해 그룹화되어 진다. 이러한 지역 연산의 합은 분류 함수 $f_A(X)$ 와 $f_B(X)$ 로 연산된다. 마지막 연산식 (11), (12)는 분류 함수를 이용하여 최종 테스트 영상이 어느 클래스에 포함 될 것인가를 보여 주고 있다.

원래 PNN알고리즘을 검색을 위한 연관성 추정 방법으로 사용할 경우 메모리개념이 없기 때문에 검색에서 이전 단계에서의 학습된 내용은 다음 단계에서 특징 값의 연관성 추정을 위해 사용되지 못하는 문제점이 있다. 그 결과로 검색 성능은 단계마다 일정하게 상승하는 것이 아니라 불안정한 형태를 보이며 때로는 오히려 성능의 저하를 가져올 수 있다. 이러한 문제점을 극복하기 위해 본 논문에서는 PNN 알고리즘이 이전의 히스토리를 고려 할 수 있도록 판별식 함수를 수정하였다.



(그림 5) 피드백을 위한 히스토리 개념

우선, 이전의 모든 피드백 동작을 히스토리에 저장 시켜 놓는다. 이때 히스토리는 5가지의 인덱스를 가지며 검색된 결과의 피드백 옵션에서 아무것도 선택하지 않았을 경우는 '0', VS는 '1', S는 '2', D는 '3', VD는 '4'의 값을 부여한다. (그림 5)에서 보는 것과 같이 각 단계에서 결과로 제시된 결과 영상 중에서 사용자가 부여한 패턴 값에 따라 각 패턴에 해당되는 인덱스에 영상의 번호가 저장되고 히스토리는 새로운 영상이 질의로 입력 될 때까지 저장된다.

만약 이전 단계에서 피드백 동작에 대해 히스토리를 유지하지 않는다면, 이전 단계에서 비유사로 선택된 영상이 다시 재검색되어 나타날 수 있음으로 반복 횟수가 증가함에도 불구하고 검색 성능은 오히려 나빠지는 결과를 나타낼 수 있다. 따라서 이전 단계의 피드백 정보를 히스토리로 유지함으로써 검색 성능을 향상시킬 수 있고 사용자의 피드백 횟수를 줄여 줄 수 있다. 이에 대한 검증은 5장의 실험 결과에서 설명하도록 한다.

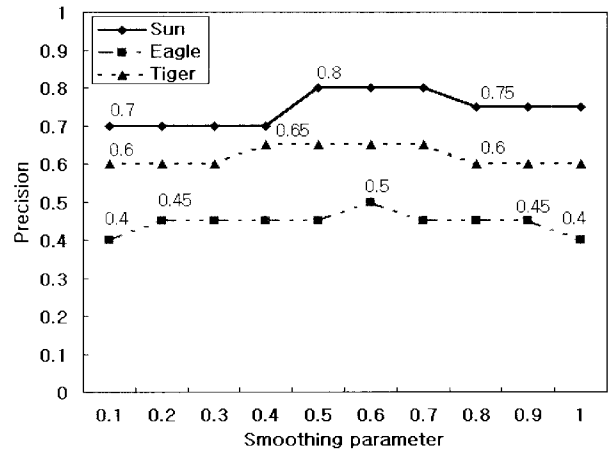
본 논문에서 제안하는 4개의 클래스를 지원하는 연관성 피드백 알고리즘은 아래와 같다.

```

Initialize j=0, for j=1,...,4; X: test pattern
//입력 및 패턴 유닛 부분
do{
    j=j+1
    netk = wj · (1-X) (Net activation) (7)
    gj = exp[(netk - 1)/σ2] (활성화 함수) (8)
    if history≡0 then gj = gj (히스토리 적용) (9)
        ≡1 then gj = gj · α
        ≡2 then gj = gj · (α/2)
        ≡3 then gj = gj/α
        ≡4 then gj = gj/α
    } while (j<4)
// summation 유닛 부분
    fA = g1 + g2 (분류 함수) (10)
    fB = g3 + g4
// 출력 유닛 부분
    {if fA ≥ fB (유사 클래스)
     then f = 1/fA (11)
     }
    {if fA ≥ fB (비 유사 클래스)
     then f = 1/fA (12)
     }

```

수식 (7)~(12)에서 j는 패턴 유닛의 수를 나타내고, g_j는 각각의 패턴 유닛에 해당하는 활성화 함수를 나타낸다. 스무딩

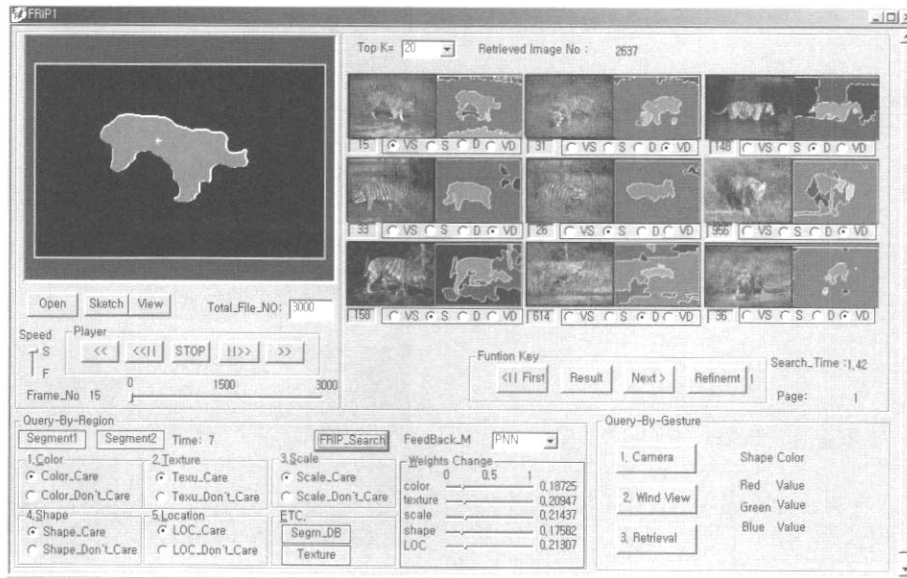


(그림 6) 스무딩 파라메타 σ의 변화에 따른 세 가지 테스트 영상의 precision

파라메타 σ는 가우시안 윈도우의 너비를 결정한다. 하지만 σ를 결정하기 위한 정형화된 방법이 없고 σ의 작은 변화는 오 분류율에 결정적인 영향을 미치지 못함으로 본 논문에서는 실험에 의해 가장 좋은 성능을 보여주는 수치를 결정하도록 하였다. 이를 위해 코렐(Corel) 영상 데이터베이스로부터 테스트 샘플로 태양, 독수리, 호랑이 영상을 선택하고 precision을 측정하여 가장 좋은 성능을 보여주는 0.6을 스무딩 파라메타 σ의 값으로 결정하였다.

수식(9)에서 α(α>1)는 히스토리를 위한 결정 파라메타로 사용자에게 의해 조절 가능한 값이다. 만약 질의 영상이 과거에 히스토리의 1과 2의 인덱스에 저장되었다면 활성 함수의 출력값은 파라메타 α에 의해 변화되어 큰 값을 갖게 되고 반대로 3 또는 4의 인덱스에 저장되어 있다면 활성 함수의 출력값은 α에 의해 작은 값을 갖게 된다. 따라서, 이전 단계에서 유사 또는 매우 유사로 선택되었던 영상은 다음 단계에서는 상위 검색 결과에 나타날 가능성이 높아진다. 수식 (10)의 f_A는 합 유닛의 분류 함수로 매우 유사(g₁)와 유사(g₂) 패턴 유닛을 위해 사용된다. f_B는 역시 비 유사(g₃)와 매우 비 유사(g₄) 패턴 유닛을 위한 분류 함수이다. 두 분류 함수에 의해 출력 유닛은 테스트 패턴 X를 '유사 클래스'와 '비 유사 클래스' 둘 중에 하나로 결정하게 된다. 즉, 수식 (11)와 (12)에서 분류 함수 f_A가 f_B보다 크다면 테스트 패턴(영상)은 유사 클래스로 선언되고 최종 결정 값 f가 재 계산된다. 반대로, 만약 분류 함수 f_B가 f_A보다 크다면 테스트 패턴(영상)은 비 유사 클래스로 선언되고 검색 결과 영상 대상에서 제외된다. 유사 클래스에서 계산된 최종 결정 값 f는 모든 데이터베이스 영상에 대해 계산되고 결정 값 f에 따라 내림차순으로 k개의 상위 영상들이 결과 영상으로 사용자에게 보여 지게 된다. 수식에서 보는 것과 같이 사용자 입력시에는 4개의 패턴을 입력 받지만 실제결과를 보여줄 때는 유사 및 매우 유사 클래스의 결과만을 사용자에게 보여주게 된다.

본 논문에서 제안하는 4단계의 연관성 피드백을 지원하는 PNN 알고리즘의 주요 과정은 다음과 같다.



(그림 7) 연관성 피드백 학습을 위해 4개의 클래스를 지원하는 FRIP 시스템 인터페이스

- (1) i 를 현재 질의영상이라 하고 특징 벡터 정규화 후에 질의 및 목표 영상간의 각 특징 값의 거리를 계산한다.
- (2) 가중치 벡터 w_j 를 모두 1로 초기화 한다.
- (3) 가중치 벡터 w_j 와 가중치기반 선형 조합 수식 (13)에 의해 질의 영상과 가장 유사한 k 개의 결과 영상을 구한다.

$$S(X, Y) = \sum_{j=1}^k w_j \cdot D(x_j, y_j) \quad (13)$$

수식 (13)에서 w_j 는 j 번째 특징을 위한 가중치를 의미하고 D 는 질의 영역의 j 번째 특징과 데이터베이스의 j 번째 영역사이의 거리를 의미한다.

- (4) 사용자는 k 개의 결과 이미지를 VS, S, D, VD 중의 하나로 선택한다.
- (5) 히스토리를 초기화 한다.
- (6) 사용자가 결과에 만족할 때까지 아래 (a)부터 (i)까지의 단계를 반복한다.
 - (a) 사용자가 선택한 k 개의 이미지를 4개의 클래스 중에서 하나로 분류하고 학습을 위해 해당 영상을 버퍼에 저장한다.
 - (b) 버퍼에 있는 학습 데이터와 수식 (4)~(6)를 이용하여 w_j 를 추정한다.
 - (c) 새 가중치 w_j 를 갱신한다.
 - (d) 새 가중치와 히스토리를 이용하는 분류 함수 결과 값을 계산한다.
 - (e) 분류 함수 결과 값을 계산하고 해당 영상이 유사 클래스와 비 유사 클래스 중 어디에 속할지 수식 (11)와 수식(12)를 이용하여 결정한다.
 - (f) (d)와 (e)의 과정을 데이터베이스의 모든 영상에 대해 실시한다.
 - (g) 유사 클래스로 분류된 영상들의 결과 값을 기준

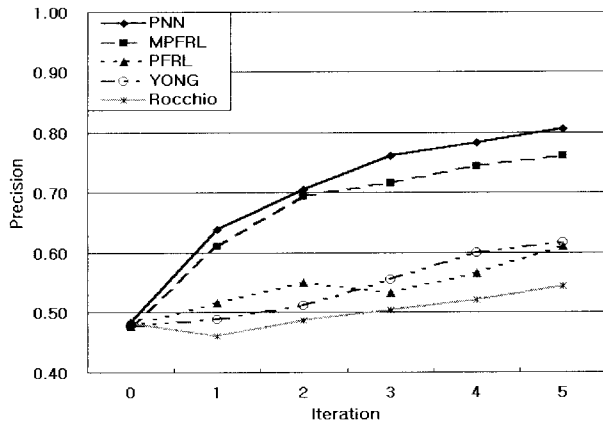
으로 내림차순으로 정렬하여 상위 k 만큼의 영상을 결과 영상으로 사용자에게 제시한다.

- (h) 사용자가 k 개의 이미지를 4개의 패턴중 하나로 선택한다.
- (i) 히스토리를 갱신한다.

5. 실험결과

FRIP 시스템은 윈도우즈 기반의 Visual C++로 개발되었다. 본 시스템에서 사용자는 브라우저를 통해 사용자가 찾길 원하는 영역이 포함된 영상을 발견하고 세그멘테이션 버튼을 눌러 분할된 영상에서 해당 영상을 더블 클릭하여 영상을 검색하도록 설계되었다. 본 시스템에 의해 분할된 영상 분할의 결과의 일부는 <http://video.kmu.ac.kr/cvpr/ROI-Resut>에서 확인할 수 있다. 또한 사용자는 보다 정확한 검색을 위해 여러 가지 특징 값을 선택적으로 사용할 수 있도록 메뉴를 구성하고 있다. 실험에서는 사용자가 모든 특징 값을 다 사용하고 상위 20개의 영상만을 결과로 제시받은 뒤 이를 통해 연관성 학습을 하도록 하였다. 상위 20개의 영상이 결과로 제시된 후에 사용자는 자신의 주관에 따라 영상을 VS, S, D, VD 중 하나로 선택하게 된다. 사용자의 선택이 끝나면 특징간의 거리 측정을 위한 5개의 가중치는 학습 과정을 통해 사용자의 인지 특성과 가깝게 재조정된다.

본 시스템의 목표는 질의 영역과 가장 유사한 영역을 포함하여 데이터베이스 영상을 검색하는 것이다. 본 실험에서는 Blobworld(<http://elib.cs.berkeley.edu/blobworld/>)[18]의 실험 방법과 유사하게 두 종류의 태양, 호랑이, 자동차, 독수리, 비행기, 교통 표지판, 꽃의 총 9개 도메인에 대해 검색 결과를 실험하였다. 실험은 코렐-포토(Corel-photo CD)로부터 무작위로 다양한 영상을 섞어놓은 3000개의 영상을 이용하였다. 모든 실험에서 검색의 성능 평가는 precision을 이용하였다.



(그림 8) 각 피드백 학습 알고리즘의 성능 평가 비교

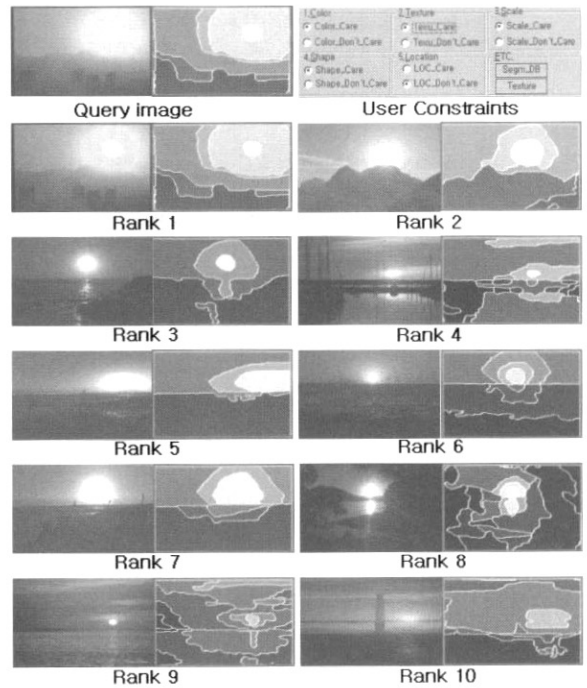
<표 1> 반복 횟수(IT)에 따른 상위 20개 영상의 검색 성능

Query	IT-0	IT-1	IT-2	IT-3	IT-4	IT-5
태양-1	0.45	0.60	0.75	0.90	0.95	1.00
독수리-1	0.45	0.75	0.75	0.90	0.90	0.95
비행기	0.20	0.50	0.60	0.60	0.65	0.70
호랑이	0.55	0.65	0.80	0.80	0.80	0.80
자동차	0.35	0.50	0.55	0.70	0.70	0.75
꽃	0.90	1.00	1.00	1.00	1.00	1.00
독수리-2	0.55	0.65	0.80	0.80	0.80	0.80
태양-2	0.45	0.70	0.75	0.80	0.80	0.80
교통 표지판	0.40	0.45	0.45	0.45	0.50	0.50
Average	0.48	0.64	0.71	0.77	0.79	0.81

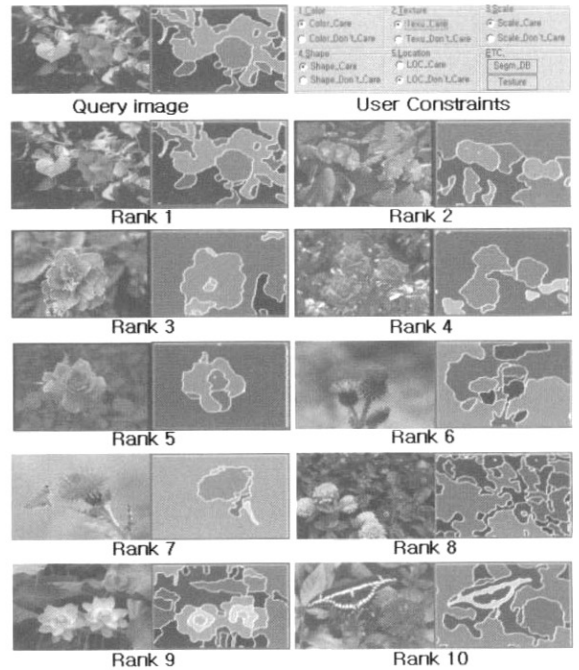
본 논문에서 제안하는 피드백 알고리즘의 성능을 검증하기 위해서 PFRL[8]과 MPFRL[4], Yong의 피드백 알고리즘[6], Rocchio의 피드백 알고리즘[7]과 성능을 비교하였다. (그림 8)은 9개 실험 영상에 대한 각 알고리즘과의 평균 검색 성능에 대한 비교 결과를 보여주고 있다. (그림 8)에서 보는 것과 같이 본 논문에서 제안한 PNN 기반 방법론은 다른 방법론과 비교하여 피드백 횟수가 증가 할수록 성능이 좋아질 뿐만 아니라 다른 방법론에서 나타나는 성능의 진동 현상이 발생하지 않는다. 다른 방법론에서 성능의 진동 현상이 발생하는 이유는 가중치만을 갱신했기 때문에 이전 단계에서 검출되었던 비유사 영상이 다시 검색되는 오류로 인한 것이다. 하지만 본 논문에서는 히스토리를 사용함으로써 이러한 문제점을 해결 할 수 있을 뿐만 아니라 사용자에게 4개의 클래스를 지원함으로써 보다 정확하고 사용자 인지(perception) 지향적인 검색 결과를 얻을 수 있다.

<표 1>은 9개 실험 영상에 대해 본 논문에서 제안하는 알고리즘을 사용하였을 경우의 검색 성능을 각 피드백 단계 별로 보여 주고 있다. 표에서 보는 것과 같이 검색 성능은 진동하지 않고 꾸준히 증가함을 알 수 있다.

<표 1>에서 보는 것과 같이 연관성 피드백 알고리즘을 사용하지 않을 경우 검색 성능은 약 48%에 불과 했다. 하지만 5번의 반복 후에 평균 검색 성능은 약 81%로 향상됨을 알 수 있다. (그림 9)부터 (그림 10)은 9개의 검색 영상 중에 3개의 질의 영상에 대한 검색 결과를 보여주고 있다.



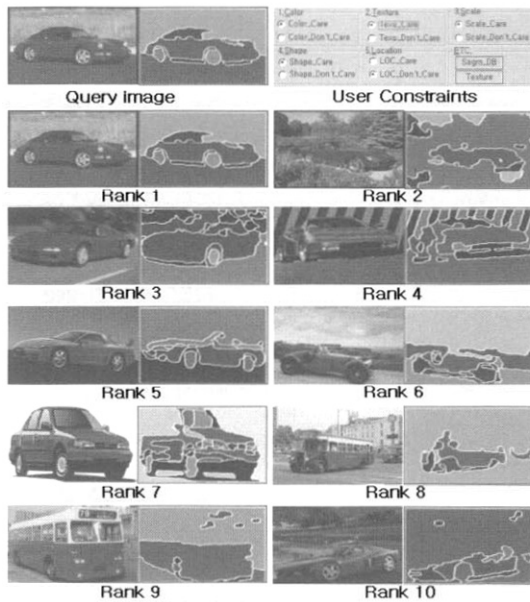
(그림 9) 태양에 대한 5단계 반복 후에 최종 검색 결과(왼쪽 : 원 영상, 오른쪽 : 세그멘테이션 영상)



(그림 10) 꽃에 대한 5단계 반복 후에 최종 검색 결과(왼쪽 : 원 영상, 오른쪽 : 세그멘테이션 영상)

6. 결 론

본 논문에서는 확률적 신경망(PNN)을 이용하여 4개의 클래스를 지원하는 연관성 피드백 알고리즘을 제안하고 있다. 본 논문의 학습 과정에서 특징 거리를 측정하기 위한 가중치는 사용자의 피드백 동작에 의해 사용자의 인지 특성과 유사하게



(그림 11) 자동차에 대한 5단계 반복 후에 최종 검색 결과(왼쪽 : 원 영상, 오른쪽 : 세그멘테이션 영상)

제조정된다. 가중치 조정 후에 분류 동작에서 가중치는 각 패턴 유닛에 적용되고 히스토리와 최종 분류 함수에 의해 상위 k개의 영상만이 사용자에게 결과로 보여 지도록 설계되었다. 본 논문의 연관성 학습 방법은 PNN을 사용하였기 때문에 특징 벡터가 독립적이어야 한다는 가정을 할 필요가 없고 시스템으로 하여금 보다 상세한 클래스의 추가를 허용하고 있다. 또한, 연관성 피드백을 위해 4개의 클래스를 지원하고 있기 때문에, 피드백 과정에서 사용자의 의도를 보다 정확하게 파악할 수 있고 사용자의 모든 피드백 동작을 저장해 놓음으로써 다음단계의 성능 향상을 보장 할 수 있다. 실험결과에서 보듯이 검색 결과는 진동하거나 감소되지 않고 반복 횟수에 따라 꾸준히 증가됨을 알 수 있다. 본 논문에서는 제안하는 피드백 학습 알고리즘을 바탕으로 향후 과제로 영상 내에서 중요 영상(OOI: Object-of-Interest)을 자동으로 추출해 내고 이를 질의 영상으로 사용하는 연구를 지속할 것이다.

참 고 문 헌

[1] M. Flickner, W. Niblack, D. Petkovic, W. Equitz and R. Barber. "Efficient and Effective Querying by Image Content," *Research Report #RJ 9203(81511)*. IBM Almaden Research Center, San Jose, 1993.

[2] M. Carson, S. Thomas, J. M. Belongie, and J. Malik. "Blobworld: A system for region-based image indexing and retrieval," *In Proc. Int. Conf. Visual Inf. Sys.*, 1999.

[3] Y. Ribner, L. J. Guibas, and C. Tomasi. "The earth mover's distance, multi dimensional scaling, and color-based image retrieval," *Proceeding of the ARPA Image Understanding Workshop*, pp.661-668, May, 1997.

[4] B.C. Ko, J. Peng, and H. Byun. "Region Based Image Retrieval

Using Probabilistic Feature Relevance Feedback," *Pattern Analysis and Application(PAA)*, Vol.4, pp.174-184, 2001.

[5] J. Peng, B. Bhanu, and S. Qing, "Probabilistic Feature Relevance Learning for Content-Based Image Retrieval," *Computer Vision and Image Understanding*, Vol.75, No.1/2, pp.150-164. 1999.

[6] Y. Rui, T. S. Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.8, No.5, pp.644-655, Sept., 1998.

[7] J.J. Rocchio, "Relevance feedback in information retrieval," In Gerard Salton, editor, *The SMART Retrieval System-Experiments in Automatic Document Processing*, pp.313-323, Prentice Hall, Englewood Cliffs, N.J., 1971.

[8] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Query Databases through Multiple Examples," *In proceeding of the 24th VLDB Conference*, New York, 1998.

[9] B. C. Ko, H. Byun, "FRIP: A Region-based Image Retrieval Tool Using Automatic Image Segmentation and Stepwise Boolean AND Matching," *IEEE Transaction on Multimedia*, Vol.7, Issue 1, pp.105-113, Feb., 2005.

[10] Y. Rui and T. Huang, "Optimizing Learning in Image Retrieval," *IEEE Int.Conference on Computer Vision and Pattern Recognition*, June, 2000.

[11] N. Vasconcelos and A. Lippman, "Bayesian Relevance Feedback for Content-Based Image Retrieval," *IEEE Workshop on Content-based Access of Image and Video Libraries*, pp.63-67, 2000.

[12] C. Meilhac and C. Naster, "Relevance Feedback and Category Search in Image Databases," *IEEE Int. Conference on Multimedia Computing and Systems*, pp.512-517, 1999.

[13] S. D. Macarthur, C. E. Brodley, and C.R. Shyu, "Relevance Feedback Decision Trees in Content Based Image Retrieval," *IEEE Workshop on Content-based Access of Image and Video Libraries*, pp.68-72, 2000.

[14] P. Hong, Q. Tian, T. S. Huang, "Incorporate support vector machines to content-based image retrieval with relevance feedback," *IEEE Int. Conference on Image Processing*, pp.750-753, 2000.

[15] D. F. Specht, "Probabilistic Neural Networks and Polynomial Adaline as Complementary Techniques for Classification," *IEEE Trans. on Neural Networks*, Vol.1, pp.111-121, March, 1990.

[16] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, A Wiley-Interscience Publication, Second Edition, 2000.

[17] L. Wu, C. Faloutsos, K. Sycara, and T. R. Payne, "FALCON: Feedback adaptive loop for content-based retrieval," *Int. Conf. Of Very Large Database(VLDB)*, pp.10-14, Sept. 2000.

[18] C. Carson, S. Belongie, H. Greenspan, J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.24 No.8, pp.1026-1038, 2002.



고 병 철

e-mail : niceko@kmu.ac.kr

1998년 경기대학교 전자계산학과(이학사)

2000년 연세대학교 대학원 컴퓨터학과
(공학석사)

2004년 연세대학교 대학원 컴퓨터학과
(공학박사)

2004년~2005년 삼성전자 통신연구소 책임연구원

2005년~현재 계명대학교 컴퓨터공학과 전임강사

관심분야: 내용기반 영상검색, 컴퓨터 비전 및 패턴인식



남 재 열

e-mail : jynam@kmu.ac.kr

1983년 경북대학교 전자공학과(공학사)

1985년 경북대학교 대학원 전자공학과
(공학석사)

1991년 University of Texas at Arlington
전기공학과(공학박사)

1985년~1987년 한국전자통신연구소 연구원

1991년~1995년 한국전자통신연구소 선임연구원

1995년~현재 계명대학교 컴퓨터공학과 부교수

관심분야: 영상압축, 영상통신, 멀티미디어 시스템